

LÉVY-SHEFFER SYSTEMS AND THE LONGSTAFF-SCHWARTZ ALGORITHM FOR AMERICAN OPTION PRICING

STEFAN GERHOLD

ABSTRACT. Glasserman and Yu (Ann. Appl. Probab. 14, 2004, p. 2090) have investigated the mean square error in the Longstaff-Schwartz algorithm for American option pricing, assuming that the underlying process is (geometric) Brownian motion. In this note we provide similar convergence results for the standard Poisson, Gamma, Pascal, and Meixner processes, pointing out the connection of the problem to the Lévy-Sheffer systems introduced by Schoutens.

1. INTRODUCTION

The least squares Monte Carlo approach by Longstaff and Schwartz [4] has become the standard method for pricing American (or Bermudan) interest rate derivatives by Monte-Carlo simulation. It proceeds by backward induction and estimates value functionals by regression on a prescribed set of basis functions. Its convergence analysis was commenced in the original paper [4], and was carried out in detail by Clément, Lamberton, and Protter [1]. In both papers it is assumed that the number of basis functions is fixed, while the number of Monte Carlo paths grows. Glasserman and Yu [3] (abbreviated GY in what follows) have analyzed settings in which these two parameters increase together. In particular, they have shown that the number of basis functions may grow only roughly logarithmically, if the underlying process is Brownian motion or geometric Brownian motion.

The present note corroborates their conjecture that their analysis should be extendable to other settings with similar results. In this way, we add some theoretical evidence that a large number of basis functions is detrimental on convergence, a plausible fact from a practical point of view [7]. On the other hand, we provide an application of the neat martingale properties that Schoutens [8] found for certain Lévy processes and families of orthogonal polynomials.

In the following section we recall Longstaff and Schwartz's algorithm and describe the problem that GY [3] treated. Our presentation is deliberately brief; see GY [3] for more on the context of the problem. Section 3 recalls the notions of Sheffer system and Lévy-Meixner system [8]. Besides Brownian motion, this theory yields four processes that lend themselves to the analysis: the standard Poisson, Gamma, Pascal, and Meixner process. In Section 4 we assume that there are only three exercise opportunities, resulting in a single regression, and show how fast the number of simulation paths must increase in order to ensure convergence of the Longstaff-Schwartz algorithm for a growing number of basis functions. Asymptotic

Date: February 14, 2008.

2000 Mathematics Subject Classification. Primary: 62P05; Secondary: 33C45.

Key words and phrases. Option pricing, Dynamic Programming, Monte Carlo, Regression, Lévy-Meixner systems.

This work has been done under the financial support by Christian Doppler Laboratory for Portfolio Risk Management (PRisMa Lab) <http://www.prismalab.at/>. The fruitful collaboration and support by the Bank Austria Creditanstalt (BA-CA) and the Austrian Federal Financing Agency (ÖBFA) through CDG is gratefully acknowledged.

inversion of this result yields rows two to five of Table 1. The last column indicates the maximal number of basis functions for which the mean square error tends to zero as the number of paths tends to infinity. Finally, Section 5 contains an analogous bound for the multi-period setting, which is weaker, but upon inversion still leads to the same critical asymptotic rate as the single-period case. In the course of the proofs, it turns out that the different critical rate of Brownian motion stems from the slower growth of the connection coefficients pertaining to the associated Lévy-Meixner system, viz. the Hermite polynomials.

Process	Basis polynomials	#Basis functions
Geometric Brownian motion	Monomials	$\sqrt{\log N}$
Standard Poisson	Charlier	$\log N / \log \log N$
Gamma	Laguerre	$\log N / \log \log N$
Pascal	Meixner	$\log N / \log \log N$
Meixner	Meixner-Pollaczek	$\log N / \log \log N$
Brownian motion	Hermite	$\log N$

TABLE 1. The highest possible number of basis functions for N paths

2. REVIEW OF THE ALGORITHM, PREVIOUS RESULTS, AND NOTATION

For the reader's convenience, our notation closely follows that of GY [3]. Suppose that our (discounted) asset follows a Markov process S_t . The American (actually Bermudan) option may be exercised at the times $0 = t_0 < \dots < t_m$. The payoff from exercise is $h_n(S_{t_n})$, for given functions h_n , $0 \leq n \leq m$. By dynamic programming, the option value at time t_0 equals $\max\{h_0(S_{t_0}), C_0^*(S_{t_0})\}$, where the continuation values C^* are given by

$$\begin{aligned} C_m^*(x) &= 0, \\ C_n^*(x) &= \mathbf{E}[\max\{h_{n+1}(S_{t_{n+1}}), C_{n+1}^*(S_{t_{n+1}})\} \mid S_{t_n} = x], \quad 0 \leq n < m. \end{aligned}$$

Longstaff and Schwartz [4] approximate these continuation values by a linear combination of basis functions ψ_{nk} :

$$C_n^*(x) \approx \sum_{k=0}^K \beta_{nk} \psi_{nk}(x) = \beta_n^T \psi_n(x),$$

where $\beta_n = (\beta_{n0}, \dots, \beta_{nK})^T$ is a vector of real numbers which is estimated by regression over the simulated paths, and $\psi_n(x) = (\psi_{n0}(x), \dots, \psi_{nK}(x))^T$.

The variant of the algorithm to be analyzed proceeds as follows. Start with the final continuation value $\hat{C}_m = 0$ and the final option value $\hat{V}_m = h_m$. For $n = m-1, \dots, 1$ generate N sample paths $\{S_{t_1}^{(i)}, \dots, S_{t_{n+1}}^{(i)}\}$, $1 \leq i \leq N$, and set

$$\begin{aligned} \hat{\gamma}_n &= \frac{1}{N} \sum_{i=1}^N \hat{V}_{n+1}(S_{t_{n+1}}^{(i)}) \psi_n(S_{t_n}^{(i)}), \\ \hat{\beta}_n &= \Psi_n^{-1} \hat{\gamma}_n, \\ \hat{C}_n &= \hat{\beta}_n^T \psi_n, \\ \hat{V}_n &= \max\{h_n, \hat{C}_n\}. \end{aligned}$$

Finally, the initial continuation value is $\hat{C}_0(S_0) = N^{-1} \sum_{i=1}^N \hat{V}_1(S_{t_1}^{(i)})$, whence the initial option value $\hat{V}_0 = \max\{h_0(S_0), \hat{C}_0(S_0)\}$ is found.

Note that the matrix

$$(1) \quad \Psi_n = \mathbf{E}[\psi_n(S_{t_n})\psi_n(S_{t_n})^T]$$

has to be estimated by its sample counterpart in practice, but is available explicitly in the examples that GY [3] and we consider. In the single-period case $m = 2$, the question that GY [3] treated is as follows. Suppose that there is a representation

$$(2) \quad h_2(S_{t_2}) = \sum_{k=0}^K a_k \psi_{2k}(S_{t_2}),$$

with unknown constants a_k . This assumption is not too restrictive; an infinite series representation of this kind has to be assumed anyways to get convergence of the algorithm, and since we are interested in $K \rightarrow \infty$, we can suppose that (2) is a good approximation of the payoff at t_2 . Furthermore, assume that the martingale property

$$(3) \quad \mathbf{E}[\psi_{2k}(S_{t_2}) \mid S_{t_1}] = \psi_{1k}(S_{t_1})$$

holds. How fast may K tend to infinity compared to N , while assuring that the mean square error of β tends to zero? To this end, they established the bounds

$$(4) \quad \sup_{|\beta|=1} \mathbf{E}[|\beta - \hat{\beta}|^2] \leq \frac{\|\Psi_1^{-1}\|^2}{N} \sum_{j=0}^K \sum_{k=0}^K \mathbf{E}[\psi_{2j}(S_{t_2})^2 \psi_{1k}(S_{t_1})^2]$$

and

$$(5) \quad \sup_{|\beta|=1} \mathbf{E}[|\beta - \hat{\beta}|^2] \geq \frac{1}{N\|\Psi_1\|^2} \sum_{k=0}^K \mathbf{E}[\psi_{2K}(S_{t_2})^2 \psi_{1k}(S_{t_1})^2] - \frac{1}{N}.$$

Here and in what follows, $|\cdot|$ denotes the Euclidean vector norm, and $\|\cdot\|$ denotes the Euclidean (or Frobenius) matrix norm. The proofs of the estimates (4) and (5) are short; the bulk of the work of GY [3] lies in the concrete examples (Brownian motion and geometric Brownian motion), and in the general analysis of the multi-period case (see Section 5).

The martingale property (3) is convenient for estimating the expectations in these two bounds. Another useful property is orthogonality of the basis functions. If S_t is Brownian motion, then GY [3] have shown that for N paths the highest K for which the mean square error tends to zero is roughly $\log N$. Hermite polynomials are natural basis functions in this case. If the underlying process is geometric Brownian motion and monomials are used as basis functions, then K may only be as high as $\sqrt{\log N}$. In this note we show that the analogous rate for the Poisson, Gamma, Pascal, and Meixner processes is in between, viz. $\log N / \log \log N$.

3. LÉVY-MEIXNER SYSTEMS

A source of basis functions and processes that satisfy martingale equalities of the type (3) are the Lévy-Meixner systems introduced by Schoutens [8]. Recall that Meixner [5] has determined all sets of orthogonal polynomials $Q_k(x)$ that satisfy Sheffer's condition

$$f(z) \exp(xu(z)) = \sum_{k=0}^{\infty} Q_k(x) \frac{z^k}{k!}$$

for some formal power series f and u with $u(0) = 0$, $u'(0) \neq 0$, and $f(0) \neq 0$. Schoutens introduces a time parameter t via

$$f(z)^t \exp(xu(z)) = \sum_{k=0}^{\infty} Q_k(x, t) \frac{z^k}{k!}$$

and shows how an infinitely divisible characteristic function, and thus a Lévy process, can be defined by f and u under appropriate conditions. Building on Meixner's characterisation, five sets of orthogonal polynomials $Q_k(X_t, t)$ and associated Lévy processes X_t are determined, which satisfy martingale equalities of the type

$$\mathbf{E}[Q_k(X_t, t) \mid X_s] = Q_k(X_s, s), \quad 0 \leq s \leq t.$$

This furnishes the connection between Sheffer (resp. Lévy-Meixner) systems and condition (3). There are five Lévy-Meixner systems, constructed from Hermite polynomials, Charlier polynomials $C_k(x, \mu)$, Laguerre polynomials $L_k^{(\alpha)}(x)$, Meixner polynomials $M_k(x; \mu, q)$, and Meixner-Pollaczek polynomials $P_k(x; \mu, \zeta)$, respectively. The resulting Lévy processes X_t are standard Brownian motion B_t , the standard Poisson process N_t , the Gamma process G_t , the Pascal process P_t , and the Meixner process H_t , respectively. See Schoutens [8] for the definitions of all these processes and families of orthogonal polynomials; we just note that the processes are characterized by the increment characteristic functions

$$\begin{aligned} \mathbf{E}[\exp(iu(B_{t+1} - B_t))] &= e^{-u^2/2}, \\ \mathbf{E}[\exp(iu(N_{t+1} - N_t))] &= \exp(e^{iu} - 1), \\ \mathbf{E}[\exp(iu(G_{t+1} - G_t))] &= \frac{1}{(1 - iu)}, \\ \mathbf{E}[\exp(iu(P_{t+1} - P_t))] &= \frac{1 - q}{1 - qe^{iu}}, \quad 0 < q < 1, \\ \mathbf{E}[\exp(iu(H_{t+1} - H_t))] &= \left(\frac{\cos(\frac{1}{2}\pi - \zeta)}{\cosh(\frac{1}{2}u - (\frac{1}{2}\pi - \zeta)i)} \right)^2, \quad 0 < \zeta < \pi. \end{aligned}$$

Brownian motion is not of interest to use, since the corresponding last line of Table 1 has been established by GY [3]. As for the remaining four processes, in the light of condition (3), the martingale relations [8]

$$\begin{aligned} \mathbf{E}[C_k(N_t, t) \mid N_s] &= \left(\frac{s}{t}\right)^k C_k(N_s, s), \\ \mathbf{E}[L_k^{(t-1)}(G_t) \mid G_s] &= L_k^{(s-1)}(G_s), \\ \mathbf{E}[M_k(P_t; t, q) \mid P_s] &= \frac{(s)_k}{(t)_k} M_k(P_s; s, q), \\ \mathbf{E}[P_k(H_t; t, \zeta) \mid H_s] &= P_k(H_s; s, \zeta), \end{aligned} \tag{6}$$

valid for $0 < s < t$, prompt us to choose the basis functions in Table 2. (Note that $(t)_k = t(t+1)\dots(t+k-1)$ is the Pochhammer symbol.) When specializing the

Process	Notation	Basis polynomials $\psi_{nk}(x)$	Parameters
Standard Poisson	N_t	$t_n^k C_k(x, t_n)$	
Gamma	G_t	$L_k^{(t_n-1)}(x)$	
Pascal	P_t	$(t_n)_k M_k(x; t_n, q)$	$0 < q < 1$
Meixner	H_t	$P_k(x; t_n, \zeta)$	$0 < \zeta < \pi$

TABLE 2. Lévy-Meixner systems

bounds (4) and (5) to our examples, we will require the orthogonality properties

$$(7) \quad \mathbf{E}[C_k(N_t, t)C_l(N_t, t)] = t^{-k} k! \delta_{kl},$$

$$(8) \quad \mathbf{E}[L_k^{(t)}(G_t)L_l^{(t)}(G_t)] = \frac{\Gamma(k+t+1)}{k!} \delta_{kl},$$

$$\mathbf{E}[M_k(P_t; t, q)M_l(P_t; t, q)] = \frac{k!}{(t)_k q^k} \delta_{kl},$$

$$\mathbf{E}[P_k(H_t; t, \zeta)P_l(H_t; t, \zeta)] = \frac{\Gamma(k+2t)}{(2 \sin \zeta)^{2t} k!} \delta_{kl},$$

as well as a way to express the squares of the basis functions as series of basis functions. We will denote by $d_{ki}(t_n)$ the *connection coefficients* in the expansion

$$(9) \quad \psi_{nk}(x)^2 = \sum_{i=0}^{2k} d_{ki}(t_n) \psi_{ni}(x).$$

Where distinction is necessary, the connection coefficients corresponding to the four families in Table 2 will be written as $d_{ki}^P(t_n)$, $d_{ki}^G(t_n)$, $d_{ki}^{Pa}(t_n)$, and $d_{ki}^M(t_n)$, respectively. The same superscripts will adorn other quantities to distinguish the four cases, viz. the Poisson, Gamma, Pascal, and Meixner process. Throughout the paper, we write c for various positive constants whose precise value is irrelevant.

4. THE SINGLE-PERIOD PROBLEM

We now state our main result about the single-period problem, where our only exercise times are $0 = t_0 < t_1 < t_2$.

Theorem 1. *Suppose $m = 2$, that S_t is a Poisson process, and that the basis functions are as in line one of Table 2. Put $(u, v) = (10, 4)$. If the number N of paths and the number K of basis functions satisfy $N \geq K^{(u+\varepsilon)K}$ for some positive ε , then*

$$\lim_{N \rightarrow \infty} \sup_{|\beta|=1} \mathbf{E}[|\beta - \hat{\beta}|^2] = 0.$$

If $N \leq K^{(v-\varepsilon)K}$, then

$$\lim_{N \rightarrow \infty} \sup_{|\beta|=1} \mathbf{E}[|\beta - \hat{\beta}|^2] = \infty.$$

For the Gamma, Pascal, and Meixner processes, with their respective basis functions from Table 2, the same holds, if (u, v) is replaced by $(8, 8)$, $(11, 7)$, and $(8, 8)$, respectively.

The announced critical rate $\log N / \log \log N$ in Table 1 then follows from the fact that the solution of $N = K^{cK}$ satisfies $K \sim c^{-1} \log N / \log \log N$, see e.g. de Bruijn [2].

Looking at (4) and (5), we begin the proof of Theorem 1 by bounding $\|\Psi_1\|$ and $\|\Psi_1^{-1}\|$, defined by (1) and Table 2.

Lemma 2. *The norms $\|\Psi_1\|$ and $\|\Psi_1^{-1}\|$ grow at most exponentially in all four cases (Poisson, Gamma, Pascal, and Meixner), except for $\|\Psi_1^P\| \leq c^K K^K$ and $\|\Psi_1^{Pa}\| \leq c^K K^{2K}$.*

Proof. For the Poisson process we find, by (7) and Stirling's formula,

$$\|\Psi_1^P\|^2 = \sum_{k=0}^K t_1^{2k} k!^2 \leq c^K K^{2K} \quad \text{and} \quad \|(\Psi_1^P)^{-1}\|^2 = \sum_{k=0}^K t_1^{-2k} k!^{-2} \leq c^K.$$

Similarly, in the Pascal case, we have

$$\|\Psi_1^{\text{Pa}}\|^2 = \sum_{k=0}^K q^{-2k} k!^2 (t_1)_k^2 \leq c^K K^{4K} \quad \text{and} \quad \|(\Psi_1^{\text{Pa}})^{-1}\|^2 = \sum_{k=0}^K q^{2k} k!^{-2} (t_1)_k^{-2} \leq c^K.$$

The Meixner process is treated by the estimates

$$\begin{aligned} c\|\Psi_1^{\text{M}}\|^2 &= \sum_{k=0}^K \frac{\Gamma(k+2t_1)^2}{k!^2} \\ &= \sum_{k=1}^K k!^{-2} \Gamma(k+2t_1 - \lfloor 2t_1 \rfloor)^2 \prod_{i=1}^{\lfloor 2t_1 \rfloor} (k+2t_1-i)^2 + O(1) \\ &\leq \sum_{k=1}^K \prod_{i=1}^{\lfloor 2t_1 \rfloor} (k+2t_1-i)^2 + O(1) = O(K^{4t_1+1}) \end{aligned}$$

and

$$\begin{aligned} c\|(\Psi_1^{\text{M}})^{-1}\|^2 &= \sum_{k=0}^K \frac{k!^2}{\Gamma(k+2t_1)^2} \\ &\leq \sum_{k=1}^K \frac{\Gamma(k+1)^2}{\Gamma(k)^2} + O(1) = \sum_{k=1}^K k^2 + O(1) = O(K^3). \end{aligned}$$

It remains to deal with the Gamma case. The parameter $t-1$ in the martingale property (6) is not quite compatible with the orthogonality relation (8) of the Laguerre polynomials. But by the formula [9]

$$L_k^{(\alpha-1)}(x) = L_k^{(\alpha)}(x) - L_{k-1}^{(\alpha)}(x)$$

we obtain

$$(10) \quad \mathbf{E}[\psi_{1k}^{\text{G}}(G_{t_1}) \psi_{1l}^{\text{G}}(G_{t_1})] = \begin{cases} -\binom{k+t_1}{k} & k = l-1 \\ \frac{2k+t_1}{k+t_1} \binom{k+t_1}{k} & k = l \\ -\binom{k+t_1-1}{k-1} & k = l+1 \\ 0 & |k-l| \geq 2 \end{cases},$$

hence Ψ_1^{G} is tridiagonal. As (10) grows only polynomially in k , it is clear that so does $\|\Psi_1^{\text{G}}\|$. As for the inverse, note that Ψ_1^{G} is diagonally dominant, so that it suffices to bound the diagonal elements of $(\Psi_1^{\text{G}})^{-1}$ (see Nabben [6, Theorem 3.1]; note that the τ_k from that theorem are all equal to 1 in our situation.) The diagonal elements e_k of $(\Psi_1^{\text{G}})^{-1}$ can be computed recursively as [6]

$$e_{KK} = \frac{K}{K+t_1} \binom{K+t_1-1}{K-1}^{-1} \leq c^K$$

and

$$e_{k-1,k-1} = \frac{k+t_1}{k} \left(\frac{2k+t_1}{k+t_1} e_{k,k} - e_{k+1,k+1} \right), \quad 1 \leq k < K.$$

A straightforward backward induction shows that this implies

$$|e_{kk}| \leq (4(t_1+1))^{K-k+1} e_{KK}, \quad 0 \leq k < K,$$

hence $\|(\Psi_1^{\text{G}})^{-1}\|$ grows at most exponentially, too. \square

We proceed to bound the fourth order moments in (4). Using (9) and the martingale relation (3), we obtain

$$\begin{aligned}
 \mathbf{E}[\psi_{2j}(S_{t_2})^2 \psi_{1k}(S_{t_1})^2] &= \mathbf{E}\left[\sum_{i=0}^{2j} d_{ji}(t_2) \psi_{2i}(S_{t_2}) \times \sum_{s=0}^{2k} d_{ks}(t_1) \psi_{1s}(S_{t_1})\right] \\
 &= \sum_{i=0}^{2j} \sum_{s=0}^{2k} d_{ji}(t_2) d_{ks}(t_1) \mathbf{E}[\mathbf{E}[\psi_{2i}(S_{t_2}) \mid S_{t_1}] \psi_{1s}(S_{t_1})] \\
 &= \sum_{i=0}^{2j} \sum_{s=0}^{2k} d_{ji}(t_2) d_{ks}(t_1) \mathbf{E}[\psi_{1i}(S_{t_1}) \psi_{1s}(S_{t_1})].
 \end{aligned}
 \tag{11}$$

The connection coefficients from the expansion (9) are well-studied objects for various families of orthogonal polynomials. See Zeng [10] for some combinatorial properties and explicit formulas. Paraphrasing one of these results [10, Corollary 2], we have

$$d_{ki}^P(t_n) = t_n^{2k-i} k!^2 i! \sum_{s \geq 0} \frac{t_n^s}{(s-k)!^2 (s-i)! (2k+i-2s)!}, \tag{12}$$

$$d_{ki}^G(t_n) = 2^{2k+i} k!^2 i! \sum_{s \geq 0} \frac{(t_n-1)_s}{4^s (s-k)!^2 (s-i)! (2k+i-2s)!}, \tag{13}$$

$$d_{ki}^{Pa}(t_n) = (1+q)^{2k+i} k!^2 i! \frac{(t_n)_k^2}{(t_n)_i} \sum_{s \geq 0} \frac{(t_n)_s (1+q)^{-2s} q^{-s}}{(s-k)!^2 (s-i)! (2k+i-2s)!}, \tag{14}$$

$$d_{ki}^M(t_n) = (-2 \cot \zeta)^{2k+i} k!^2 i! \sum_{s \geq 0} \frac{(t_n)_s (1 + (\cot \zeta)^{-2})^s}{4^s (s-k)!^2 (s-i)! (2k+i-2s)!}. \tag{15}$$

Here it is understood that $1/n! = 0$ for n a negative integer, as is natural when extending the factorial by the Gamma function. Therefore, the sums in (12)–(15) run from $s = \max\{i, k\}$ to $s = k + \lfloor i/2 \rfloor$.

Moment bounds in the Poisson case. By (7), (11), and (12), the sum on the right-hand side of (4) can be estimated by

$$\sum_{j=0}^K \sum_{k=0}^K \mathbf{E}[\psi_{2j}^P(S_{t_2})^2 \psi_{1k}^P(S_{t_1})^2] \leq c^K \sum_{j=0}^K \sum_{k=0}^K \sum_{i=0}^{2 \min\{k,j\}} i! \left(\sum_{s \geq 0} b_{jis}^P \right) \left(\sum_{s \geq 0} b_{kis}^P \right), \tag{16}$$

where

$$b_{kis}^P := \frac{k!^2 i!}{(s-k)!^2 (s-i)! (2k+i-2s)!}.$$

It is easy to see that $b_{k+1,i,k+l+1}^P / b_{k,i,k+l}^P > 1$ for $i \geq 1$, $0 \leq l \leq i/2$, and $k \geq i-l$, hence $b_{k,i,k+l}^P$ increases in k under these conditions. From this we deduce that the s -sums in (16) increase in j respectively k :

$$\begin{aligned}
 \sum_{s=\max\{i,k\}}^{k+\lfloor i/2 \rfloor} b_{kis}^P &= \sum_{l=\max\{i-k,0\}}^{\lfloor i/2 \rfloor} b_{k,i,k+l}^P \\
 &\leq \sum_{l=\max\{i-k,0\}}^{\lfloor i/2 \rfloor} b_{k+1,i,k+l+1}^P \\
 &= \sum_{s=\max\{i,k\}+1}^{k+\lfloor i/2 \rfloor+1} b_{k+1,i,s}^P \leq \sum_{s=\max\{i,k+1\}}^{k+\lfloor i/2 \rfloor+1} b_{k+1,i,s}^P.
 \end{aligned}$$

Using this in (16) yields (recall that c may change its value in each occurrence)

$$(17) \quad \sum_{j=0}^K \sum_{k=0}^K \mathbf{E}[\psi_{2j}^{\mathbf{P}}(S_{t_2})^2 \psi_{1k}^{\mathbf{P}}(S_{t_1})^2] \\ \leq c^K K!^4 \sum_{i=0}^{2K} \left(\sum_{s=K}^{K+\lfloor i/2 \rfloor} \frac{i!^{3/2}}{(s-K)!^2 (s-i)! (2K+i-2s)!} \right)^2.$$

It is easy to see that the summand increases in i for $K \geq 0$, $0 \leq i \leq K$ and $K \leq s \leq K + i/2$. Hence we find that the portion $\sum_{i=0}^K$ of the i -sum in (17) can be bounded from above by

$$(18) \quad (K+1)K!^3 \left(\sum_{s=K}^{\lfloor 3K/2 \rfloor} \frac{1}{(s-K)!^3 (3K-2s)!} \right)^2 \leq c^K K^{5K}.$$

The last equality holds because the summand in (18) is unimodal with mode at $s = K + K^{2/3} - \frac{4}{3}K^{1/3} + O(1)$. The remaining part of the i -sum in (17) can be estimated by

$$(19) \quad \sum_{i=K+1}^{2K} i! \left(\sum_{s=i}^{K+\lfloor i/2 \rfloor} \frac{i!s!}{(s-K)!^2 (s-i)! (2K+i-2s)!} \right)^2 \\ \leq c^K \sum_{i=K+1}^{2K} i! \left(\frac{i!(K+\lfloor i/2 \rfloor)!}{\lfloor i/2 \rfloor!^2 (K+\lfloor i/2 \rfloor-i)! (i-2\lfloor i/2 \rfloor)!} \right)^2 \\ \leq c^K \sum_{i=K+1}^{2K} \frac{i!(K+\lfloor i/2 \rfloor)!^2}{(K+\lfloor i/2 \rfloor-i)!^2} \leq c^K K^{6K}.$$

Note that in the first line we have introduced the new factor $s!$ in the numerator. This makes the summand increasing w.r.t. the substitution $i \rightarrow i+1$, $s \rightarrow s+1$. Hence it suffices to keep only the summands of the s -sum with $s = K + \lfloor i/2 \rfloor$, which shows the first inequality. As for the second inequality, note that the factor $i!/\lfloor i/2 \rfloor!^2$ of the summand grows only exponentially, and that the factor $(i-2\lfloor i/2 \rfloor)!$ in the denominator is clearly negligible. Finally, the last sum in (19) has increasing summands, which together with Stirling's formula implies the last inequality. By (17), the estimates (18) and (19) show that

$$\sum_{j=0}^K \sum_{k=0}^K \mathbf{E}[\psi_{2j}^{\mathbf{P}}(S_{t_2})^2 \psi_{1k}^{\mathbf{P}}(S_{t_1})^2] \leq c^K K^{10K}.$$

In the light of (4) and Lemma 2, the value $u = 10$ for the Poisson Process in Theorem 1 is established.

As for the second assertion about the Poisson process in Theorem 1, note that, from (11),

$$\mathbf{E}[\psi_{2K}(S_{t_2})^2 \psi_{1K}(S_{t_1})^2] = \sum_{i=0}^{2K} \sum_{s=0}^{2K} d_{Ki}(t_2) d_{Ks}(t_1) \mathbf{E}[\psi_{1i}(S_{t_1}) \psi_{1s}(S_{t_1})].$$

The orthogonality property (7) and formula (12) yield

$$\begin{aligned} \sum_{k=0}^K \mathbf{E}[\psi_{2K}^P(S_{t_2})^2 \psi_{1k}^P(S_{t_1})^2] &\geq c^K \sum_{k=0}^K \sum_{i=0}^{2k} d_{Ki}^P(t_2) d_{ki}^P(t_1) i! \\ &\geq c^K d_{K,2K}^P(t_2) d_{K,2K}^P(t_1) (2K)! \\ &\geq c^K (2K)!^3 \geq c^K K^{6K}. \end{aligned}$$

The second inequality follows from retaining only the summand $k = K$, $i = 2K$. This makes the sum in (12) collapse to the summand $s = 2K$, whence the third inequality. Appealing to (5) and Lemma 2 completes the proof of the Poisson part of Theorem 1. Note that the preceding estimates can presumably be improved. This seems not worthwhile, though; since our estimate of $\|\Psi_1^P\|$ in Lemma 2 is sharp, we will not obtain equal values $u = v$ in Theorem 1 anyways, unless at least one of the bounds (4) and (5) were improved, too.

Moment bounds in the Meixner case. The proofs in the remaining three cases are very similar to the Poisson case. In the Meixner case, we have (20)

$$\sum_{j=0}^K \sum_{k=0}^K \mathbf{E}[\psi_{2j}^M(S_{t_2})^2 \psi_{1k}^M(S_{t_1})^2] \leq c^K \sum_{j=0}^K \sum_{k=0}^K \sum_{i=0}^{2\min\{k,j\}} i!^2 \left(\sum_{s \geq 0} b_{jis}^M \right) \left(\sum_{s \geq 0} b_{kis}^M \right),$$

where

$$b_{kis}^M := \frac{k!^2 s!}{(s-k)!^2 (s-i)! (2k+i-2s)!}.$$

Again, $b_{k,i,k+l}^M$ increases in k , and the remaining steps to show the upper bound are completely analogous to the Poisson case. This time the numerator factor $s!$ in the analogue of (19) appears naturally, and is not introduced artificially to force some monotonicity. Moreover, the lower bound uses the same summands as in the Poisson case. Both resulting bounds are of the form $c^K K^{8K}$, whence $u = v = 8$ in Theorem 1.

Moment bounds in the Pascal case. We can reuse the values b_{kis}^M and the estimate that we just sketched:

$$\begin{aligned} \sum_{j=0}^K \sum_{k=0}^K \mathbf{E}[\psi_{2j}^{\text{Pa}}(S_{t_2})^2 \psi_{1k}^{\text{Pa}}(S_{t_1})^2] &\leq c^K \sum_{j=0}^K \sum_{k=0}^K \sum_{i=0}^{2\min\{k,j\}} i!^3 k! \left(\sum_{s \geq 0} b_{jis}^M \right) \left(\sum_{s \geq 0} b_{kis}^M \right) \\ &\leq c^K K! (2K)! \sum_{j=0}^K \sum_{k=0}^K \sum_{i=0}^{2\min\{k,j\}} i!^2 \left(\sum_{s \geq 0} b_{jis}^M \right) \left(\sum_{s \geq 0} b_{kis}^M \right) \\ &\leq c^K K! (2K)! K^{8K} \leq c^K K^{11K}. \end{aligned}$$

The lower bound poses no new difficulties either.

Moment bounds in the Gamma case. This part is only slightly more involved. Due to (10), we have three i -sums instead of one in the analogue of (17). The right-hand side of (10) can be replaced by c^K in each of these. Then one of the three i -sums equals the i -sum in (20), and the other two differ only in an index shift $b_{k,i \pm 1,s}^M$, which can be easily bounded by polynomial factors. Thus the resulting growth rate is $c^K K^{8K}$, as for the Meixner case. The proof of Theorem 1 is complete.

5. THE MULTI-PERIOD PROBLEM

The following result shows that the critical growth rate $\log N / \log \log N$ from Table 1 carries over to the multi-period setting. We assume that a representation

analogous to (2) holds at time t_m , and that the payoff function does not grow too fast in the following sense.

Theorem 3. *Suppose that the payoff functions satisfy the growth constraint*

$$\mathbf{E}[h_n(S_{t_n})^4] \leq \max_{\nu} \left(\frac{t_{\nu+1}}{t_{\nu}} \right)^{2K} \max_{\nu,k} \mathbf{E}[\psi_{\nu k}(S_{t_{\nu}})^4], \quad 0 \leq n \leq m.$$

Then

$$\sup_{|\beta_{m-1}|=1} \mathbf{E}[|\beta_n - \hat{\beta}_n|^2] \leq N^{-1} c^K K^{(m-n+1)uK}, \quad 1 \leq n < m,$$

where u takes on the same values as in Theorem 1, that is, 10, 8, 11, 8 for S_t the standard Poisson, Gamma, Pascal, and Meixner process, respectively.

Proof. By results of GY [3] (Theorem 3 and the last formula before (18) on p. 2096) and Jensen's inequality, we have

$$\begin{aligned} \sup_{|\beta_{m-1}|=1} \mathbf{E}[|\beta_n - \hat{\beta}_n|^2] &\leq \frac{c^K}{N} \max_{1 \leq \nu < m} \|\Psi_{\nu}^{-1}\|^3 \max_{\nu,k} \mathbf{E}[\psi_{\nu k}(S_{t_{\nu}})^4]^{m-n} \max_{\nu,k} \mathbf{E}[\psi_{\nu k}(S_{t_{\nu}})^2]^2 \\ &\leq \frac{c^K}{N} \max_{1 \leq \nu < m} \|\Psi_{\nu}^{-1}\|^3 \max_{\nu,k} \mathbf{E}[\psi_{\nu k}(S_{t_{\nu}})^4]^{m-n+1}. \end{aligned}$$

Note that GY [3] assume that the moments $\mathbf{E}[\psi_{nk}(S_{t_{\nu}})^2]$ and $\mathbf{E}[\psi_{nk}(S_{t_{\nu}})^4]$ are increasing in n and k , and formulate their Theorem 3 with $\mathbf{E}[\psi_{mk}^{2(4)}]$ instead of $\max_{\nu,k} \mathbf{E}[\psi_{\nu k}^{2(4)}]$. But an inspection of their proof quickly shows that taking the max in the above estimate gets rid of the monotonicity assumption. Now note that $\|\Psi_{\nu}^{-1}\| \leq c^K$ in all our four cases by Lemma 2, and that

$$\max_{\nu,k} \mathbf{E}[\psi_{\nu k}(S_{t_{\nu}})^4] \leq \max_{\nu} \sum_{j=0}^K \sum_{k=0}^K \mathbf{E}[\psi_{\nu j}(S_{t_{\nu}})^2 \psi_{\nu k}(S_{t_{\nu}})^2] \leq c^K K^{uK},$$

where the double sum has been estimated in the proof of Theorem 1. \square

Acknowledgement. I thank Friedrich Hubalek for helpful comments.

REFERENCES

- [1] E. CLÉMENT, D. LAMBERTON, AND P. PROTTER, *An analysis of a least squares regression method for American option pricing*, Finance Stoch., 6 (2002), pp. 449–471.
- [2] N. G. DE BRUIJN, *Asymptotic methods in analysis*, Bibliotheca Mathematica. Vol. 4, North-Holland Publishing Co., Amsterdam, 1958.
- [3] P. GLASSERMAN AND B. YU, *Number of paths versus number of basis functions in American option pricing*, Ann. Appl. Probab., 14 (2004), pp. 2090–2119.
- [4] F. A. LONGSTAFF AND E. S. SCHWARTZ, *Valuing American options by simulation: A simple least-squares approach*, Rev. Financial Stud., 14 (2001), pp. 113–148.
- [5] J. MEIXNER, *Orthogonale Polynomsysteme mit einer besonderen Gestalt der erzeugenden Funktion*, J. London Math. Soc., 9 (1934), pp. 6–13.
- [6] R. NABBEN, *Two-sided bounds on the inverses of diagonally dominant tridiagonal matrices*, Linear Algebra Appl., 287 (1999), pp. 289–305. Special issue celebrating the 60th birthday of Ludwig Elsner.
- [7] V. PITERBARG, *Pricing and hedging callable LIBOR exotics in forward LIBOR models*, Journal of Computational Finance, 8 (2004/05), pp. 65–119.
- [8] W. SCHOUTENS, *Stochastic processes and orthogonal polynomials*, vol. 146 of Lecture Notes in Statistics, Springer-Verlag, New York, 2000.
- [9] G. SZEGÖ, *Orthogonal polynomials*, vol. XXIII of Colloquium Publications, American Mathematical Society, fourth ed., 1975.
- [10] J. ZENG, *Weighted derangements and the linearization coefficients of orthogonal Sheffer polynomials*, Proc. London Math. Soc. (3), 65 (1992), pp. 1–22.

CHRISTIAN DOPPLER LABORATORY FOR PORTFOLIO RISK MANAGEMENT (PRISMA), VIENNA
UNIVERSITY OF TECHNOLOGY, WIEDNER HAUPTSTRASSE 8-10, A-1040 VIENNA, AUSTRIA
E-mail address: `sgerhold at fam.tuwien.ac.at`